

Recommending peers for learning: Matching on dissimilarity in interpretations to provoke breakdown

Citation for published version (APA):

Rajagopal, K., Van Bruggen, J., & Sloep, P. (2017). Recommending peers for learning: Matching on dissimilarity in interpretations to provoke breakdown. *British Journal of Educational Technology*, 48(2), 385-406.
<https://doi.org/10.1111/bjet.12366>

DOI:

[10.1111/bjet.12366](https://doi.org/10.1111/bjet.12366)

Document status and date:

Published: 01/03/2017

Document Version:

Early version, also known as pre-print

Document license:

CC BY-NC-ND

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 06 May. 2023

Open Universiteit
www.ou.nl



THIS IS A PREPRINT OF: Rajagopal, K., van Bruggen, J. M. and Sloep, P. B. (2015), Recommending peers for learning: Matching on dissimilarity in interpretations to provoke breakdown. British Journal of Educational Technology. doi: 10.1111/bjet.12366

Recommending peers for learning: matching on dissimilarity in interpretations to provoke breakdown

Kamakshi Rajagopal

Jan M. van Bruggen

Peter B. Sloep

Dr. Kamakshi Rajagopal, Dr. Jan M. van Bruggen and Dr. Peter B. Sloep work at the Welten Institute of the Open University of the Netherlands. Their common research interests include networked learning (specifically but not exclusively for professionals) and the application of language technologies in educational technology. Correspondence address: Open Universiteit, Valkenburgerweg 177, 6419 AT Heerlen. email: Kamakshi.rajagopal@gmail.com

Abstract

People recommenders are a widespread feature of social networking sites and educational social learning platforms alike. However, when these systems are used to extend learners' Personal Learning Networks, they often fall short of providing recommendations of learning value to their users. This article proposes a design of a people recommender based on content-based user profiles, and a matching method based on dissimilarity therein. It presents the results of an experiment conducted with curators of the content curation site Scoop.it!, where curators rated personalized recommendations for contacts. The study showed that matching dissimilarity of interpretations of shared interests is more successful in providing positive experiences of breakdown for the curator than is matching on similarity. The main conclusion of this article is that people recommenders should aim to trigger constructive experiences of breakdown for their users, as the prospect and potential of such experiences encourages learners to connect to their recommended peers.

Practitioner notes

What is already known about this topic

- Educational practitioners create personal networks to promote learning.
- Automated people recommenders on social networking sites look for people who are similar to you – not for people who are different from you.
- By curating content, lifelong learners can contextualize the topics of their interest.
- “Breakdown”, a well-known theoretical concept where the encounter of something unexpected makes a person reconsider their previously held thoughts, is useful to understand learning but as yet elusive in its technological application.

What this paper adds

- A matching algorithm that operationalizes breakdown by latching onto dissimilarity between learners proves to offer more value to a learner than do similarity algorithms.
- A learner who may find particular recommended content generally relevant, is likely to experience breakdown with respect to some specific aspects of this content.
- A learner who has experienced breakdown due to the content someone else has shared, is likely to want to contact that person.

Implications for practice and/or policy

- Further developing the notion of dissimilarity helps to identify people’s specific understanding of a topic better.
- Learners and educators using people recommenders to improve learning should evaluate the performance of automated recommendations on relevance for learning.

Introduction

Personal Learning Networks (PLNs) are online social networks that are used to support the continuous professional development of lifelong learners in a personalized manner (Van Harmelen, 2008; Granovetter, 1983; Rajagopal, Joosten-ten Brinke, Van Bruggen & Sloep, 2012). Like most online social networks, these networks are dense with people and resources with which its members connect online and offline (Reich, Subrahmanyam & Espinoza, 2012; Brandtzaeg, 2012). Consequently, one of the most common forms of personalized support in PLNs concerns navigation through the network, to efficiently and effectively offer learners the most relevant content or contacts. A popular form of personalized navigation support is recommender systems i.e. systems that filter out the most relevant content or fellow-networkers for an individual user, based on their previous activities or existing connections in the network (Resnick & Varian, 1997; Manouselis et al., 2012; Fazeli, Drachsler, Brouns & Sloep, 2012). When recommender systems focus on connecting people in the network with each other, they are called social matching systems or people recommenders (Resnick & Varian, 1997; Terveen & McDonald, 2005). All recommender systems conceptually consist of three components (i) a user profile to characterise an individual user, (ii) a matching algorithm to determine which users should be recommended to each other, and (iii) a user interface, to introduce a user to their recommended match. The aim of a people recommender is to introduce a learner to a relevant contact, based on certain previously determined criteria. Of these, the matching algorithm is the interest of this article.

Most recommender systems are based on the principle of similarity: they are designed to seek out people or resources that have traits similar to a user's background or to things that the user has an interest in (Resnick & Varian, 1997). This principle is embedded into learning technology through the algorithms used. Content-based matching algorithms find content that is similar to content that the user has already studied. Collaborative filtering matching algorithms look for items used by others who have used items similar to the ones used by the target user. Even social network-based matching starts from similarity: the assumption is that if user A has several ties in common with an unconnected user B, the odds are that user A also knows and wants to connect with user B. The similarity lies in the promotion of common ties with others (Liu & Lee, 2010). Features of matching algorithms can be combined to create hybrid systems, which combine different matching strategies to achieve results that are more relevant to the user (Burke, 2002; Sie, Drachsler, Bitter-Rijkema & Sloep, 2012). Research shows that content-similarity matching algorithms perform better at recommending new contacts than social network-based recommenders (Chen, Geyer, Dugan, Muller & Guy, 2009). The principle of similarity in recommendation is equally widespread in technology-enhanced learning (Drachsler, Verbert & Duval, 2013).

Crucially, though, the principle of similarity is flawed when it comes to connecting people for learning purposes. Beyond a general similarity of interest, learners who seek to learn from other people, in general, look for someone who can give a specific, different perspective on a topic (Rajagopal, Verjans, Costa & Sloep, 2012). They look for content or perspectives that are unlike their own. Learners pursue instances where such a mismatch of views can be experienced (Jenkins, 2009). This stands in stark contrast with, for example, movie or book recommenders, where a viewer or reader is interested in more of roughly the same.

In the literature, the experience of mismatch between a learner's current understanding and the elements challenging this understanding is called 'breakdown'. We find this term in literature on independent learners and learning professionals (Koschmann, Kuutti & Hickman, 1998, Winograd & Flores, 1986; Schön 1983) and in literature relating to organisational learning (Weick, Sutcliffe & Obstfeld, 2005). Breakdown presumes that learners hold certain mental views about a particular topic or process (termed *personal understanding* by Stahl (2006)). A useful definition of the term breakdown is the "disruption in the normal functioning of things, forcing the individual to adopt a more reflective or deliberative stance toward ongoing activity" (Koschmann, Kuutti & Hickman, 1998, p26). Breakdown can be triggered through many factors and in many environments, such as through a comment or pertinent questions in a conversation with another person, through the offering of new, relevant content, or through unexpected behaviour from others. When breakdown occurs, a learner's received understanding appears to be inadequate and consequently, she starts re-assessing her current thoughts. If and when a learner follows up such breakdown with discussions with other learners to (re-)establish common ground and to develop a shared understanding, the learner enters a social process of interaction, shared understanding and knowledge-building (Stahl, 2006). In this way, the learner engages in collaborative knowledge-building and sense making (Clark & Brennan, 1991; Dubberly & Pangaro, 2009; Dillenbourg, 1999). Parenthetically, the experience of breakdown alone does not necessarily constitute a learning experience. Breakdown and the subsequent knowledge-building require an investigative learning attitude, the will to investigate and question one's own cherished assumptions. Lacking such an attitude, breakdown will become merely identifying a difference of perspective at best, or a discussion at cross-purposes at worst, not necessarily leading to further positive, shared knowledge-building.

When looking at the design of people recommenders for learning purposes then, this notion of breakdown points to the need for a matching algorithm based on a different principle, namely that of *dissimilarity*. Similarity at a *general level* may be necessary to establish a new contact, but dissimilarity on *specifics* would then trigger the interest to connect. Matching on dissimilarity means that, once common interests between learners A and B have been established, A and B focus on finding specific meaningful differences between them. 'Meaningful' in this context refers to their need to agree on parts of their understanding, but also the need to be different in a significant way so as

to potentially experience breakdown and consequent refinement of understanding for one or preferably both of them.

The first hypothesis of this article looks at the principle underlying people recommenders based on content matching, i.e. recommenders that use content as a proxy to profile the learner: we hypothesise that when a learner deems generally relevant content that is shared by another learner, the first learner is likely to have an experience of constructive breakdown due to encountering specific mismatches that occur between her own understanding and the other's understanding of that content. Consequently, an experience of breakdown leads to the desire to connect with the now recommended person. This is illustrated in Figure 5.1. Regarding the matching principle, we surmise that selecting recommended contacts through their content by matching on specific dissimilarities rather than by matching on general similarity is more likely to create an experience of breakdown and a subsequent desire to connect.

HYPOTHESIS I: People recommenders with content-based user profiles and matching methods work on the following model of recommendation:

1. Recommended content that is deemed generally relevant indirectly leads to users wanting to connect with each other. *We will see this reflected in a high correlation score between the estimation of the relevance of recommended content and the desire to connect.*
2. Content that is deemed relevant by the user creates the possibility of an experience of constructive breakdown for the user. *We will see this reflected in a high correlation score between a user's relevance score for recommended content after use and the extent to which users indicate that they have experienced constructive breakdown, i.e. have re-assessed their thoughts on a topic because of the recommended content.*
3. A user who experiences breakdown, provoked in relation to some recommended content deemed relevant, will want to connect with the user who has provided this content. *We will see this reflected in a high correlation score between the report of an experience of breakdown and an indication of a desire to connect with the recommended user, through a discussion.*

<<< INSERT FIGURE 5.1 ABOUT HERE >>>

A matching strategy for dissimilarity for people recommenders starting from content can be operationalised in several ways. We will look at methods capable of revealing meaningful differences between users' understandings. Following theoretical concepts from the field of knowledge building closely, where the mind is seen as an environment in which concepts are connected, disconnected and reconnected continuously (Bereiter, 2004), we take the words a user employs to express her thoughts as a means to gain insight into how she relates the concepts that the words

jointly describe. The most frequently used expressions of thoughts in online environments are self-authored, short, written texts in natural language. So, suitable resources can be platforms where users are actively writing such texts, on their understanding of a topic.

For this experiment, we selected the content curation website Scoop.it! as it is one of the most widely used and well-known curation websites, primarily based on text and conducive to lifelong learning. Scoop.it! users, or *curators*, can manage a page dedicated to a topic of their choice and select and gather relevant resources on this topic. They do this by creating a *scoop*, i.e. bookmarking and sharing a link to the resource they consider valuable. Moreover, the curator may (but is not obliged to) expand on the content of this scooped resource by providing some meta-level insights or comments, in the form of added text, thus contextualising the content of the resource. This is especially interesting as this gives us insight into how the curator positions the topic of the curated link, with regard to other links and/or other topics.

So a Scoop.it! curator is not merely sharing links, but contextualising these links within a unique individual perspective, thereby becoming an actively sharing learner. This cognitive nature of the activity of curation leads us to think that Scoop.it! curators have a learning attitude and actively engage with their topic of interest. By interpreting a curator's shared resources together with the contextualizing, short, self-authored texts on Scoop.it! as the curator's personal user profile, we can use the profile as a starting point for people recommenders.

Scoop.it! allows its curators much freedom in how they use the platform, thus creating huge differences between curators: some post many scoops, others few only; some write context with every scoop they share, others give context for some selected posts only; some are very elaborate in the accompanying text, others add just few contextualizing words.

The short written texts on Scoop.it! can be reduced to sets of words using natural language processing (NLP) techniques, by extracting lexically important words from the learner-authored texts. Although these techniques are technically heavier than e.g. using platforms with curator-given tags, there is an added potential, which we will illustrate with an example; in it the lexical words have been italicised. Imagine the following situation:

Person A's comment: "this *article shows* that *human intelligence* is *related* to *upbringing*"

Person B's comment: "we *see* that *human intelligence depends* on *natural talent* and *upbringing*"

In this example, person B is talking about more concepts than person A. But more importantly, the collection of lexical words indicates that person B talks of a different (more nuanced) type of relation between the concepts (namely, *depends* vs. *related*). This subtle but semantically significant difference cannot be captured through the use

of curator-given tags alone, as tags are primarily used to describe something at a general level.

The use of NLP in recommendation has so far been limited to conversational recommender systems, where the system engages in a dialogue with the user to narrow down the search space (e.g. critiquing, elaboration) (Grasch & Felfernig, 2015). We will use NLP to create an informative content-based user profile of the curators within our people recommender. As these user profiles stem directly from each curator’s texts and vocabulary use, we may assume that they are close to their understanding of the topic. This follows similar trends in recommendation where text-based user profiles are made by mining user-written reviews, to augment other approaches to matching users with products (Chen & Chen, 2015). This approach is as yet unexplored in people recommendation (Chen, Chen & Wang, 2015). The profiles can then be matched according to general similarity and specific dissimilarity:

Word-space similarity: We can measure to what extent two learners are talking about the same concepts, which we define as words with a common stem. A word-space is a conceptual space, which is defined by all the words used together in a particular context. To measure the similarity between two instances within that conceptual space, we can calculate how many words they share. In this experiment, we group all the keywords a learner uses in her collection of authored texts. The user profile of a learner then consists of one large set of keywords, which may be compared with another learner’s set, against the backdrop of all the keywords used by all the learners in the group. This type of similarity has been used before, often using Jaccard similarity (where only the presence or absence of an item is relevant) or cosine similarity (where the degree of presence of an item is relevant). In this research, we have used Jaccard similarity.

Tagset dissimilarity: We can measure to what extent learners agree with each other in the relations they see between concepts. We assume that the co-occurrence of two lexically important words within one learner-authored text snippet is an indication that the learner sees a direct relation between the concepts described by the words. We do not necessarily know in which way the learner relates the concepts, but we know that the learner perceives some relation between the concepts since they occur in the same text snippet. This is plausible as the learner-authored text snippets we use are limited in length, thereby increasing the likelihood of a curator-perceived relation between the concepts.

A simple method to achieve this measurement is to group together the keywords emerging from one learner-authored text snippet, thereby creating separate collections of keywords for a particular individual. The content and the contextualization added by the curator can be described in terms of “tags”, the term we use to describe lexical words with a unique semantic value. A collection of tags, which we call a “tagset”, indicates that the curator sees some relation between the semantic values of the tags in the tagset. In other words, the tags in that tagset have some semantic or pragmatic relation, which has been imprinted on them by the curator. Each tagset can be traced back to the combination of a resource and a contextualizing authored text, which creates a richer, more layered user profile. Consequently, the learner is then

represented by several, relatively small tagsets, instead of by one large set of tags. Tagsets can be created for all authored text snippets of a learner. Tagsets of different curators can be compared in several ways. For our purposes, we look for curators who have partially overlapping tagsets. This would indicate, that as a pair they both see relations between two or more concepts but individually they also relate these concepts to other concepts in ways different to each other.

<<< INSERT FIGURE 5.2 ABOUT HERE >>>

Figure 5.2 illustrates the two types of user representations that we will use: on the left, one set of concepts per learner for Jaccard similarity and on the right, several sets of tagsets containing concepts per learner for tagset dissimilarity. As both these methods start from the same types of user profiles (keywords extracted from learner-authored text snippets) and work with similar concepts of matching (overlap in sets), we expect that the resulting matches will be similar for most curators. However, as the first method emphasizes similarity between learners and the second emphasizes dissimilarity, it is expected that both will bring up differences in their relevance to and appreciation by the curator. This brings us to our second and last hypothesis:

HYPOTHESIS 2: Matching a person with another person based on principles of similarity in the general use of concepts and matching based on dissimilarity in the interpretation and relationship of specific concepts results in different user perceptions.

1. Matching users on the basis of similarity in their general use of concepts will identify more common relevant content to users than matching them on the basis of dissimilarity in the interpretation of concepts.
2. Matching users on the basis of dissimilarity in the specific interpretation of a relationship between concepts will give users the perception of a higher experience of breakdown than matching them on the basis of similarity in the general use of concepts.
3. The more tags are shared between users, the more the recommended content will be deemed relevant, and the more experiences of breakdown will be reported.

Methods

Participants

The participants of our experiment were Scoop.it! curators maintaining at least one topic page, and adhering to the following criteria:

- The page is categorized under at least one of the following keywords used in the Scoop.it! general search: “Higher Education”, “Networked Learning”, “Educational” or “Learning”
- The topic page is curated by an individual and not by a group of people or an organization.

- The topic page has a Scoop.it! score above 35 (based on the number of independent views).
- The primary language of the page is English. This means most posts are English posts and most curator-added insights are in English.
- The data on the topic page was publically available.

After the aggregation and anonymisation of all personal data (primarily the names of contributors) 240 Scoop.it! topic pages, each associated with its unique curator, were left for further analysis. (List1) From these 240 pages, we retained only those with curator-authored insights or meta-reasoning texts, creating a subset of 148 Scoop.it! topic pages (List2). Reducing this subset to concepts and collections of tags and grouping pages with the same curator gave us a list of 139 curators (List3). When asked to fill in a personalised survey to evaluate recommended matches, we received filled-in usable surveys from 46 curators (List4).

Instruments and Measures

We created two instruments to collect measurements for this experiment: a people recommender system and a personalised set of probes. Figure 5.3. describes how the instruments are related to each other.

<<< INSERT FIGURE 5.3. ABOUT HERE >>>

I. Recommender system

The people recommender consists of two matching algorithms, with their related user profiles.

Content-based User profile. The user database contains two types of content-based user profiles for each of the 139 participants: (i) a set of tags W for the Jaccard matching and, (ii) a set of tagsets S for the tagset matching.

These user profiles were created from the 240 collected Scoop.it! pages. The last 200 individual scoops ('posts') of each of the pages were gathered, and only those posts with a curator-added meta-reasoning in the form of an insight were retained. Natural language processing tools were then applied to transform the posts to tags and tagsets

- The text of each post was tokenized using the Natural Language Toolkit (NLTK in Python) creating token-tagsets, where each word in the curator-created post is considered to be an instance of a lexeme (or unique unit of lexical meaning), and as such, an element of a set (illustrated in Figure 5.4).
- The most frequent words (more than 150 occurrences; generally words such as pronouns, common verbs, etc.) and the least frequent words (fewer than 3 occurrences) were removed from the token-tagsets in order to retain lexically relevant words.
- The tokens were subsequently stemmed to relate words with common stems, such as e.g. "collaborate" and "collaboration" (Porter stemmer

in the NLTK) and duplicates were removed to keep unique tags over the whole dataset. Posts from topic pages belonging to the same curator were grouped.

<<< INSERT FIGURE 5.4 ABOUT HERE >>>

Table 5.1. shows some numerical data covering the changes in the database following these steps. The data shows that every Scoop.it! post with curator-added context introduces just fewer than 1 new concept, or in other words, nearly every post with curator-added context introduces a new concept.

<<< INSERT TABLE 5.1. ABOUT HERE >>>

Matching algorithms. We used two matching algorithms for calculating recommendations.

1. Jaccard similarity

Using set-of-tags user profiles, we calculated similarity between two curators based on their common usage of tags, using the well-known Jaccard coefficient. This coefficient looks at the proportion of instances where both curators use the same tag against the sum of the instances where a tag is used by one curator and not the other and the instances where a tag is used by both curators. The following formulation expresses the Jaccard similarity coefficient s :

Where Q is the set of tags used by curator A, and D is the set of tags used by curator B,

$$s = \frac{|Q \cap D|}{|Q \cup D|}$$

or

$$s = \frac{p}{p + q + r}$$

where

p = number of tags that are present in both Q and D

q = number of tags that are present in Q but not D

r = number of tags that are present in D but not Q

$|Q \cap D|$ = cardinality of intersection between Q and D

$|Q \cup D|$ = cardinality of union between Q and D

Jaccard similarity coefficients were calculated between all curators, resulting in a distance matrix containing curator-to-curator distance measure for each curator pair in the database. For each curator in the database, we ranked the matches according to

their matching scores establishing a best ranked Jaccard match and a worst ranked Jaccard match.

2. Tagset dissimilarity

Using the set-of-tagsets user profiles, this algorithm matches two curators based on how they use tags together. We distinguish four types of possible overlaps between two arbitrary tagsets T and S : (i) the tagsets are identical $T=S$ (TypeA), (ii) one tagset is the subset of the other $T \subset S$ (TypeB), (iii) the two tagsets partially overlap in that they share some elements but also both have their own elements s and $t \rightarrow (T \cap S = \emptyset)$ AND $\exists s, t: s \in S \ \& \ t \in T \ \& \ s \notin T \ \& \ t \notin S$ (TypeC); (iv) the two tagsets do not share any elements $T \cap S = \emptyset$ (TypeD).

Of these, the distribution of TypeC overlaps shows the extent to which two curators are talking about the same topics, but each relates these common topics to other topics. This form of overlap brings to the fore critical differences in understanding between two curators.

We created a distance matrix with the distance measure, $\sigma(A, B)$, to match two curators (A and B), on the basis of the proportion of TypeC tagset overlaps between them and the 10-base logarithm of the number of matched tagsets for the pair A and B , to take into account the disparity in the number of tagsets per participant. A larger number of matched tagsets (and subsequently a higher log) indicates more evidence for the match. This results in the formulation of a matching score σ , which is a measure of the extent to which two curators are similar in their discussion on general shared interests but dissimilar in the specific interpretation of those interests:

$$\sigma(A,B) = \frac{\text{\#Type C}}{a * b} * \log(a*b)$$

where:

$a = |S(A)|$ = number of tagsets of person A

$b = |S(B)|$ = number of tagsets of person B

\#TypeC : the number of Type C overlaps between A and

For each curator in the database, we ranked the matches according to their matching scores establishing a best-ranked tagset match and a worst-ranked tagset match.

We calculated a total of 7,404,820 matched tagset pairs between the 139 participants, which are distributed across the four types of overlap as in Table 5.2. The vast majority of matched tagset pairs have no overlap (TypeD, 79,3%), giving a clear indication that most Scoop.it! posts are unique in the topics they cover, and curators are fairly specific in how they connect various topics. TypeC (partial overlap) count for 20,6%, whereas full overlap (TypeA) and subset (TypeB) hardly occur (0.1%).

<<< INSERT TABLE 5.2 ABOUT HERE >>>

Each unique matched pair of participants has two matching measures, $s(A, B)$ (Jaccard) and $\sigma(A, B)$ (tagset). Table 5.3 shows a summary of the distributions of these scores, and also indicates the strong correlation between them.

<<< INSERT TABLE 5.3 ABOUT HERE >>>

The outcome of the recommender system is two best-ranked recommendations (one based on Jaccard similarity and one on tagset dissimilarity) and two worst-ranked recommendations (one based on Jaccard similarity and one on tagset dissimilarity) for each curator in the database. The best-ranked matches aim to present content that is highly likely to be perceived as relevant (i.e. sufficiently different and interesting) by the curator, whereas the worst-ranked matches present content that is least likely to be deemed relevant by the curator.

The use of partial overlap between matched tagsets (TypeC) is a solid indicator for the relationship between two participants' understanding of a topic. The tagset scores have a greater variance than the Jaccard scores. The high correlation ($r=0.72$) between the Jaccard scores and the tagset score shows that both matching methods uncover similar factors from the comparisons. However, there are differences between the methods.

II. Probes

The second instrument was a set of personalized probes per participant to evaluate the recommendations made for them in our experiment. Through three 5-point Likert-scale questions, we wanted to gauge to what extent the recommendations fulfill the following needs for an individual curator:

- Content relevance: This feed contains new and relevant information for me.
- Experience of breakdown: The comments and opinions expressed in this feed make me re-assess my thoughts about this topic.
- Desire to connect: I would like to engage in a discussion with the curator of this feed.

Note: In what follows, variables that measure each of these constructs are italicized. Thus *content relevance* measures the relevance that a curator ascribes to a particular piece of content; etc.

Research design

After identifying the best-ranked and worst-ranked matches between the participants, according to the methods of similarity and dissimilarity, we evaluated the calculated matches by presenting them to the participants in personalized sets of probes. Each set

of personalised probe contained a set of four anonymised Scoop.it! feeds selected for each curator on the basis of the results of the matching.

T1: One best-ranked tagset match

T2: One best-ranked Jaccard match

T3: One worst-ranked tagset match

T4: One worst-ranked Jaccard match

Each feed had three randomly selected posts from the original aggregation.

The 139 curators from our database were invited to evaluate their four recommended matched persons on the variables of *content relevance*, *experience of breakdown* and *desire to connect* through an online form. Participants were asked expressly to confirm that we were allowed to use the data from their probes for analysis. 46 completed surveys were returned.

We have the following variables associated with each matched pair of participants and associated with each participant (Table 5.4).

<<< INSERT TABLE 5.4 ABOUT HERE >>>

Figures 5.5 and 5.6 show a graphical representation of the hypotheses and the expected results. For the first hypothesis, correlation scores are calculated to establish the internal consistency between the variables. We expect highly correlated user scores at both ends of the Jaccard and tagset ranking.

<<< INSERT FIGURE 5.5 ABOUT HERE >>>

For the second hypothesis, we compare the user scores for all three variables between the two matching methods as follows: the scores of best-ranked dissimilarity matches are compared to the scores of best-ranked similarity matches; the scores of worst-ranked dissimilarity matches are compared to the scores of worst-ranked similarity matches.

<<< INSERT FIGURE 5.6. ABOUT HERE >>>

Results

In the analyses, we adopted the common practice of treating Likert scales as interval scales.

For our first hypothesis, we expected a positive correlation between the three variables: *content relevance*, *experience of breakdown* and *desire to connect*, according to both methods of comparison (general similarity and specific dissimilarity). Table 5.5 shows the correlations between *content relevance* and *experience of breakdown* for the four groups of matches.

<<< INSERT TABLE 5.5 ABOUT HERE >>>

Content relevance is positively associated with the *experience of breakdown* as is evidenced by the correlations of a medium-to-high strength. We also see that the correlation coefficients for best dissimilarity and worst dissimilarity are both about the same figure (0.5). In other words, the variability of the data remains relatively constant for both the best and the worst-ranked matches. A much bigger variability is observed in similarity-based matching, as the correlation is strong (0.79) for the worst-ranked matches but low (0.35) for the best-ranked matches. So, the curator evaluations of the best similarity matches for these correlations are much more diverse.

This difference in variability is also noticeable in the correlations between *experience of breakdown* and *desire to connect*. Table 5.6 shows these figures and the partial correlation controlling for *content relevance*.

<<< INSERT TABLE 5.6 ABOUT HERE >>>

The partial correlation scores are high in dissimilarity-based matching, both for the best-ranked (0.69) and worst-ranked matches (0.49). In similarity-based matching, the worst-ranked matches show a strong (partial) correlation, whereas there is much more variability in the best-ranked matches. The results are illustrated in Figure 5.7.

<<< INSERT FIGURE 5.7. ABOUT HERE >>>

The lack of variability in dissimilarity matching can be seen as an indicator of the reliability of this matching method. These results seem to indicate that matching on dissimilarity gives us a much more stable method to predict a curator's experience of content and desire to connect with the recommended match. Matching on similarity performs well in the worst-ranked matches, where there is none to very little similarity. Evaluations of best similarity matches are however much more varied.

To investigate the second hypothesis on the differences in user evaluation between the two matching methods, we conducted within-subjects repeated measures ANOVAs for the three variables (*content relevance*, *experience of breakdown* and *desire to connect*) for the best-ranked matches (T1 and T2) and for the worst-ranked matches (T3 and T4), considering two conditions: tagset dissimilarity matching and Jaccard similarity matching. Table 5.7 and table 5.8. show an overview of all the statistics.

Best-ranked Matches. As we only compare the variables in two conditions, Mauchly's test showed that the assumption of sphericity is not violated. The analysis showed a main effect of the matching method in *content relevance*, but the results are not significant at the 5%-level. However, note they would be if we were to use an

only slightly more relaxed alpha. Post-hoc comparisons between the two conditions showed a significant difference in mean between dissimilarity matching and similarity matching ($p=0.017$). This suggests that dissimilarity-based recommendations are evaluated higher on *content relevance* than similarity-based recommendations. With a medium effect size and a low observed power, more research is needed to arrive at reliable conclusions. For *experience of breakdown*, the analysis showed a significant effect of the matching method and a significant difference in the mean of the user evaluations ($p=0.017$). This shows that the best-ranked dissimilarity-based matches have been evaluated significantly higher than the best-ranked similarity-based matches. The large effect size and the relatively high observed power make it reasonable to conclude the difference observed is significant. There was no significant effect of the matching method for *desire to connect* and the low effect size and the low observed power indicate a need for further research to determine the relation between matching on dissimilarity and the desire to connect. In conclusion, the results for the best-ranked matches indicate that the tagset dissimilarity matching method seems to predict the curator evaluations for *experience of breakdown* better than the Jaccard similarity matching method.

<<< INSERT TABLE 5.7. ABOUT HERE >>>

<<< INSERT TABLE 5.8. ABOUT HERE >>>

Worst-ranked Matches. As we only compared the variables in two conditions, Mauchly's test showed that the assumption of sphericity is not violated. There was a non-significant effect of the matching method observed for *content relevance* and the low effect size and the observed power show inconclusive results, indicating the need for more research. The analysis showed a significant main effect of the matching method for *experience of breakdown*. Post-hoc comparisons between conditions showed a significant difference in the mean between dissimilarity matching and similarity matching ($p=0.010$). This indicates that the recommendations of worst-ranked dissimilarity-based matches were evaluated significantly lower than the similarity-based recommendations. The large effect size and high observed power make it reasonable to infer a significant difference. The analysis also showed a significant effect of the matching method for *desire to connect*. Post-hoc comparisons between conditions showed a significant mean difference between the dissimilarity and similarity matching methods ($p=0.011$). This too shows that dissimilarity-based recommendations are scored lower than similarity-based recommendations for worst-ranked matches. The medium-large effect size and the observed power here support this significant difference. In conclusion, the tagset dissimilarity matching method seems to predict the curator evaluations of the variables *experience of breakdown* and *desire to connect* better.

All results for hypothesis 2 are illustrated in Figure 5.8.

<<< INSERT FIGURE 5.8. ABOUT HERE >>>

Further research is needed to assess if this performance of the matching methods can be captured in a measure for matching performance to qualify the matching methods. For the last part of the second hypothesis, we investigated whether the variable *maximum intersection* size (i.e. the maximum number of tags that two participants share) is correlated with *content relevance*, *experience of breakdown* or *desire to connect*. Although no significant results came out of the correlation analysis, plotting the maximum intersection size of a match against the related evaluation scores of the match does give us some insight into their relationship. Figure 5.9 shows six heat maps. Heat maps graphically illustrate the values in colour intensity – the darker an area on a heat map, the more people have chosen a particular evaluation score. To enhance the clarity of the plots, we grouped the two lowest scores of the Likert scale and the two highest scores of the Likert scale, effectively creating a 3-point scale. The figure shows the original 5-point scales on the left and the reduced 3-point scales on the right.

The *content relevance* heat maps show a consistent positive evaluation as the maximum intersection size increases, as shown by the darker top right corner. This suggests that more shared tags in tagsets affect the perceived relevance of content. The *experience of breakdown* (and to a lesser extent *desire to connect*) has most answers situated on the left side of the heat maps. In other words, a low maximum intersection size, seemingly with a cut-off point of three shared tags, seems to evoke either positive or negative reactions from the curators for their perceived experiences of breakdown. More shared tags in tagsets do not necessarily affect the perceived experience of breakdown or the desire to connect. Further investigation is needed to verify these initial findings.

<<< INSERT FIGURE 5.9 ABOUT HERE >>>

Discussion

We assume that the Scoop.it! curators who participated in this research have a learning attitude and as such, are open to experiences of breakdown and seeking (actively or reactively) to connect with others on the platform. Also we keep in mind that the recommendations calculated are not symmetrical, i.e. person A's top match may be person B, but person B's top match is not necessarily person A. As a consequence, for any matched pair one-directional or unbalanced relationships are likely to exist.

The results support the first hypothesis. The curator evaluation of the probes shows that when curators perceive an experience of breakdown through the feed of a matched person, they also exhibit a desire to connect with that person, as suggested by the strong correlation between the two variables. The probes' results also indicate that providing content that is deemed relevant contributes to the experience of breakdown. However, the correlation scores are generally lower here, perhaps indicating that there are other factors that contribute to the experience of breakdown

apart from content relevance. The results strongly support that a curator's experience of breakdown with respect to one particular person, will contribute to a desire to connect with that person. Experience of breakdown can therefore be a good starting point to create new connections in a learner's personal learning network. *Content relevance* is a variable that can be manipulated to trigger experiences of breakdown, opening up new avenues for designed learning experiences. More research is needed to identify possible other factors that contribute to a curator's experience of breakdown.

The correlation analysis indicates that dissimilarity-based matching gives us a stable method to predict a curator's evaluation of content relevance and their experience of breakdown, related to their desire to connect with the recommended match. Matching on similarity performs well for the worst-ranked matches, but the curator evaluations of best similarity-based matches show a greater variation.

The second hypothesis is also partially confirmed by the results. There are differences between the curator evaluations of dissimilarity matches and similarity matches, which is primarily noticeable for the variable *experience of breakdown*: best-ranked dissimilarity-based matches are rated higher on experience of breakdown than best-ranked similarity-based matches. Similarly, worst-ranked dissimilarity-based matches are rated lower than similarity-based matches. A similar effect is noticeable for the variable *desire to connect* in the worst-ranked matches. The data do not show significant differences between dissimilarity-based matches and similarity-based matches for *content relevance*. These results together with the results of the correlation analysis show that the method of tagset dissimilarity is a more suitable method for identifying suitable contacts for a particular curator, with the expectation that breakdown can occur in an interaction between them. However, as there is no significant difference in the evaluation of content relevance, it seems appropriate to conclude that experiences of breakdown depend on more factors than content alone, in line with the conclusions on the first hypothesis. Based on these results, the dissimilarity-based matching method seems to predict the experience of breakdown better, but further research should determine where the strengths of this method lie.

Additionally, there is a relation between the maximum number of tags a curator shares with a match and the rating that the curator gives on the variables *content relevance*, *experience of breakdown* and *desire to connect*. The plotted results seem to indicate a critical value of three shared tags to trigger a (strong or weak) experience of breakdown. Higher numbers of shared tags do not necessarily evoke higher scores in the experience of breakdown. However, higher numbers of shared tags do seem to result in higher ratings of content relevance. This result points to the need for strategies other than mere common usage of tags for triggering experiences of breakdown in curators. For example, combining tagset dissimilarity with qualitative filtering of the tags could possibly improve the performance of the dissimilarity algorithm presented in this article.

The technical approach presented in this article elicits meaningful words from a user's own social media posts. Recommendation through sentiment analysis uses a similar text-mining approach, taking into account the positive or negative sentiment

expressed by the user towards an individual or an object (Alahmadi & Zeng, 2015). The contribution of dissimilarity matching (as operationalized in this article) lies in that it makes no statement about the user's positive or negative stance towards a potential match, but only looks at the degree to which the user is talking about the same concepts as a potential match (i.e. through partial overlap in tagsets), which is the most interesting for learning purposes.

The people recommender system presented in this article addresses many of the design issues discussed in recently published work by Mayer, Jones and Hiltz (2015). They identify that mobile social matching systems need to address the user's relational, social and personal contexts appropriately to successfully introduce strangers to each other. To achieve this, they propose several new design concepts including *contextual rarity* and *contextual oddity* to refer to situations where two potential users are more likely to connect with each other because they share a common attribute, or in fact, do not share a common attribute within a particular social context. The strategy of dissimilarity-based matching is a first step in the operationalization of these two design concepts, as it successfully elicits those users with meaningful or interesting differences of interpretation against a common shared interest or topic. The work presented in this article can be a first answer to the "open challenges" that the new design concepts of Mayor, Jones and Hiltz set to designers of people recommenders (Mayer, Jones & Hiltz, 2015).

Conclusions

Our results should affect the recommendation strategies used by people recommenders for learning, which currently primarily aim to recommend others who are in some way similar to the curator. The results of this experiment suggest that a curator is likely to connect with another person who has been responsible for triggering a experience of breakdown caused by provision of relevant content, or other (still to be researched) factors. Therefore, people recommenders would benefit from aiming to create interactions between curators where they can experience breakdown. This is in line with previous research in educational technology design that focuses on creating interaction between learners (Fetter, Berlanga, Sloep, Van der Vegt, Rajagopal & Brouns, 2012; Sloep, 2013). A better understanding of which factors influence experiences of breakdown, and how these factors can be manipulated, will allow the design of better learner support tools.

Our second conclusion concerns the matching methods used in people recommenders. This study has shown that using a dissimilarity-based matching method may be more successful in predicting curator evaluation of perceived experience of breakdown than similarity-based matching. It also seems to be a good strategy to predict curator evaluations of perceived content relevance and the expressed desire to connect with a recommended curator. Further research needs to assess if this performance of the matching methods can be captured in numerical measures.

This study has also shown some avenues for improving the tagset dissimilarity matching method. Initial observations show that there seems to be a cut-off point at about three shared tags between the curator and the recommended match: fewer

triggers an experience of breakdown, more not necessarily. Monitoring which tags are shared and which are not shared between two people may allow for better matching. In particular, such monitoring can be implemented using linguistic principles. The underlying assumption is that the types of concepts that are shared between learners gives a better insight into their understanding of the domain, which in turn allows one to make a better representation of them in automated systems. For example, the use of domain-specific marker words could give some insight into which subsection of the domain they might be situated in. Likewise, the use of a syntactically and semantically annotated lexicon and the use of a grammar might provide more elaborate descriptions of the curator's posts and consequently the curator's intended meanings. This in turn might create opportunities for calculating the dissimilarity between curators in more detail, potentially increasing chances of triggering breakdown. In short, dissimilarity-based matching is an important matching method for educational people recommenders, deserving further investigation.

Limitations

The experiment is limited in a number of ways, potentially affecting the scope of the conclusions drawn.

1. We selected the curation website Scoop.it! as the platform of our experiment. From the experiment, we learnt that this particular platform is often used as a personal content management platform, rather than as a place to connect with others. The respondents' feedback clearly showed that curators use different platforms for different purposes, and many respondents did not connect with others on Scoop.it! on principle. Therefore, positioning the people recommender of this article as a recommendation service within Scoop.it! may have been a factor in how the respondents answered the probes. However, this is a general consideration regarding social networking sites: when does a curator go beyond the content shared by a person to the person herself?
2. Another consequence of our choice of Scoop.it! as a platform is that this is a directed network, i.e. a network where the connections between the curators' topic pages is not reciprocal. One curator follows another, but this connection is not necessarily mutual. This is similar to the social network of Twitter, but unlike other popular social networks, such as LinkedIn and Facebook. Using the matching strategies described in this article on such reciprocal networks may bring up other results.
3. The comparison of the two matching methods in terms of correlation of the scores has its limits. A more qualitative approach needs to be taken better to understand the differences between them, for which more sophisticated measuring instruments are necessary. The similarity measure used in this article is the Jaccard similarity coefficient, a well-known base method. However, recent research work has explored the extension of similarity-based methods with network data (Fazeli, Loni, Drachsler & Sloep, 2014). The dissimilarity method presented in this article also needs to be compared to these more sophisticated approaches to recommendation.

4. Although we started with a target group of 240 Scoop.it! curators, the number of participants who chose to fill in the curator-evaluation surveys was limited to 46 only. There were several reasons for this:

- We reached the participants primarily through social media (i.e. Twitter and Scoop.it!), as this was the only direct contact information we had of many of the participants. This method of working had advantages (namely, immediacy of answers, triggering enthusiasm of participants) but also quite a few limitations (if we did not get a response, it was not always clear if this was because they ignored the invitation or if they had not seen it).
- Some participants showed initial interest, but did not follow through to complete the survey.
- The invitations relied purely on the goodwill of the participants to join in. We did not offer any incentives apart from insight into which Scoop.it! feeds were recommended for them, after they had filled in the survey.

5. It is always possible that the people recommended to the curator are not deemed relevant or interesting to connect with since the curator already has knowledge about the different opinion brought in and does not agree with it. The final decision is left to the curating learner. The experience of breakdown is only partially determined by content relevance, other factors, yet to be determined, may also play a role.

Future Work

There are several avenues for future work following this research.

First is a further exploration of the benefits of curation websites for profiling learners. There is high-quality information in these websites that reveals the interests, opinions and understandings of a learner, which presumably could be used to create better learner support in networked learning contexts.

A second, more immediate avenue is a further development of the tagset dissimilarity algorithm and method. Further research needs to refine the matching method to exploit its strengths and make it more reliable in predicting the curator's evaluation of the recommendation. One could imagine the development of a language-specific lexicon that awards weights to the use of certain marker concepts, a domain-specific lexicon that awards weights to the use of certain topic concepts, or a lexicon that is tied to an ontology such as Wordnet.

A third avenue is to understand if the performance of the matching methods based on similarity and dissimilarity can be captured in a single numerical measure, which would then qualify the matching method in terms of its performance.

A fourth avenue is the exploration of technological support for sensemaking and the experience of breakdown, not only from a collective perspective, but from the individual learners' contribution as well. There has been extensive research on how technology can be used to support sensemaking processes for teams and groups of people. This technology has largely been concerned with facilitating the dialogue and conversations between group members (Conklin, 2006), and augmenting them with

various types of related contextual information (Brown, Downie & Buckingham Shum, 2012; De Liddo & Buckingham Shum, 2010). The emphasis of these technologies is on the collaborative space. However, sensemaking is now happening on a global scale through various current social networking technologies (e.g. LinkedIn Groups, Facebook pages, Twitter groups, MOOCs, etc). Unlike the sensemaking technologies that focus on the collaborative space, these networking technologies focus on the role of the individual learner within the collective. There are a lot of opportunities to make previously learnt lessons concerning collaborative sensemaking fit into current circumstances.

REFERENCES

- Alahmadi, D. H., & Zeng, X. J. (2015). ISTS: Implicit social trust and sentiment based approach to recommender systems. *Expert Systems with Applications*.
- Bereiter, C. (2004). Education and mind in the knowledge age. *British Journal of Educational Psychology*, 74, 127-138.
- Brandtzaeg, P. B. (2012). Social Networking Sites: Their Users and Social Implications - A Longitudinal Study. *Journal of Computer-Mediated Communication*, 17(4), 467–488. doi:10.1111/j.1083-6101.2012.01580.x
- Brown, M., Downie, a., & Buckingham Shum, S. (2012). Mapping the Matrix: Using Compendium as a Tool for Recording the Analytic Group. *Group Analysis*, 45(1), 99–115.
- Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.
- Chen, L., Chen, G., & Wang, F. (2015). Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction*, 25(2), 99-154.
- Chen, G., & Chen, L. (2015). Augmenting service recommender systems by incorporating contextual opinions from user reviews. *User Modeling and User-Adapted Interaction*, 1-35.
- Chen, J., Geyer, W., Dugan, C., Muller, M., & Guy, I. (2009, April). Make new friends, but keep the old: recommending people on social networking sites. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 201-210). ACM.
- Clark, H.H. & Brennan, S.E. (1991). Grounding in communication. In Resnick, L. B.; Levine, J. M.; Teasley, J. S. D., *Perspectives on socially shared cognition*, American Psychological Association, ISBN 1-55798-376-3
- Conklin, J. (2006a). *Dialogue mapping: Building shared understanding of wicked problems*. New York: Wiley.
- De Liddo, A. & Buckingham Shum, S. (2010). Cohere: A prototype for contested collective intelligence. In: *ACM Computer Supported Cooperative Work (CSCW 2010) - Workshop: Collective Intelligence In Organizations - Toward a Research Agenda*, February 6-10, 2010, Savannah, Georgia, USA.

- Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (Vol. 1, pp. 1–15). Oxford: Elsevier.
- Drachsler, H., Verbert, K., & Duval, E. (2013). *Recommender systems for learning* (pp. 1-20). New York, NY: Springer.
- Dubberly, H., & Pangaro, P. (2009). What is conversation? How can we design for effective conversation? *Interactions magazine*, 1–9. Retrieved from <http://www.dubberly.com/articles/page/3>
- Fazeli, S., Drachsler, H., Brouns, F., & Sloep, P. B. (2012). A Trust-based Social Recommender for Teachers. In N. Manouselis, H. Drachsler, K. Verbert, & O. C. Santos (Eds.), *2nd Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2012) in conjunction with the 7th European Conference on Technology Enhanced Learning (EC-TEL 2012), September, 18-19, 2012, Saarbrücken, Germany* (pp. 49–60). Saarbrücken (Germany). Retrieved from <http://dspace.ou.nl/handle/1820/4428>
- Fazeli, S., Loni, B., Drachsler, H., & Sloep, P. B. (2014). Which Recommender System Can Best Fit Social Learning Platforms? In C. Rensing, S. de Freitas, T. Ley, & P. Muñoz-Merino (Eds.), *Open Learning and Teaching in Educational Communities. Proceedings of the 9th European Conference on Technology Enhanced Learning (EC-TEL2014), Lecture Notes in Computer Science 8719* (pp. 84–97). Graz, Austria: Springer International Publishing.
- Fetter, S., Berlanga, A. J., Sloep, P. B., Van der Vegt, W., Rajagopal, K., & Brouns, F. (2012). Using Peer-Support to Connect Learning Network Participants to Each Other: An Interdisciplinary Approach. *International Journal of Learning Technology*, 7(4), 378–399. doi:10.1504/IJLT.2012.052212
- Granovetter, M. (1983). The strength of weak ties: a network theory revisited, *Sociological theory*, 1, 201–233.
- Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. MacArthur Foundation.
- Grasch, P., & Felfernig, A. (2015). On the Importance of Subtext in Recommender Systems. In *icom* (Vol. 14, No. 1, pp. 41-52).
- Koschmann, T., Kuutti, K. & Hickman, L. (1998) The Concept of Breakdown in Heidegger, Leont'ev, and Dewey and Its Implications for Education, Mind, Culture, and Activity, 5:1,25-41, DOI: 10.1207/s15327884mca0501_3
- Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. *Expert Systems with Applications*, 37(7), 4772–4778. doi:10.1016/j.eswa.2009.12.061
- Manouselis, N., Drachsler, H., Verbert, K., & Santos, O. C. (Eds.) (2012). *Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2012)*. Published by CEUR Workshop Proceedings, 2012, Vol. 896.
- Mayer, J. M., Jones, Q., & Hiltz, S. R. (2015). Identifying Opportunities for Valuable Encounters: Toward Context-Aware Social Matching Systems. *ACM Transactions on Information Systems (TOIS)*, 34(1), 1.

- Rajagopal, K., Verjans, S., Costa, C., & Sloep, P. B. (2012). People in Personal Learning Networks: Analysing their Characteristics and Identifying Suitable Tools. In V. Hodgson, C. Jones, M. de Laat, D. McConnell, T. Ryberg, & P. Sloep (Eds.), *Proceedings of the Eighth International Conference on Networked Learning 2012* (pp. 252-259). April, 2-4, 2012, Maastricht, The Netherlands
- Rajagopal, K., Joosten-ten Brinke, D., Van Bruggen, J., & Sloep, P. B. (2012). Understanding personal learning networks: Their structure, content and the networking skills to optimally use them. . *First Monday*, 17(1), 1–12. Retrieved from <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3559/3131>
- Reich, S. M., Subrahmanyam, K., & Espinoza, G. (2012). Friending, IMing, and hanging out face-to-face: overlap in adolescents' online and offline social networks. *Developmental psychology*, 48(2), 356–68. doi:10.1037/a0026980
- Resnick, P., & Varian, H. R. (1997). Recommender Systems. *Communications of the ACM*, 40(3), 56–58.
- Schön, D.A. (1990). *Educating the Reflective Practitioner*. San Francisco (Calif.), USA: Jossey-Bass.
- Sie, R. L. L., Drachsler, H., Bitter-Rijkema, M., & Sloep, P. B. (2012). To whom and why should I connect? Co-author Recommendation based on Powerful and Similar Peers. *International Journal of Technology Enhanced Learning (IJTEL)*, 4(1), 121-137.
- Stahl, G. (2006). *Group cognition* (Vol. 106). Cambridge, MA: MIT Press.
- Terveen, L., & McDonald, D. W. (2005). Social matching: A framework and research agenda. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(3), 434.
- Van Harmelen, M. (2008). Design trajectories: four experiments in PLE implementation. *Interactive Learning Environments*, Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/10494820701772686>
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. *Organization Science*, 16(4), 409–421.
- Winograd, T. A., & Flores, C. F. (1986). *Understanding computers and cognition: A new foundation for design*. Intellect Books.

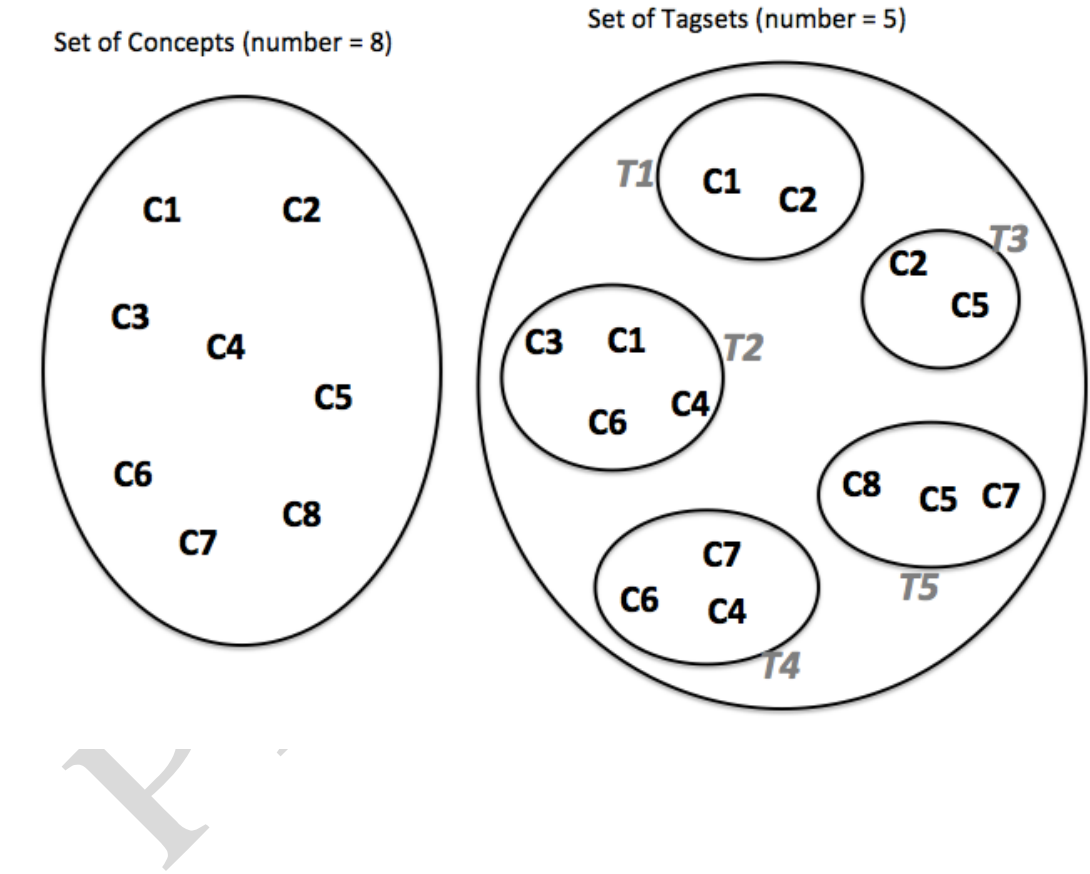
LIST of FIGURES AND TABLES

Figures

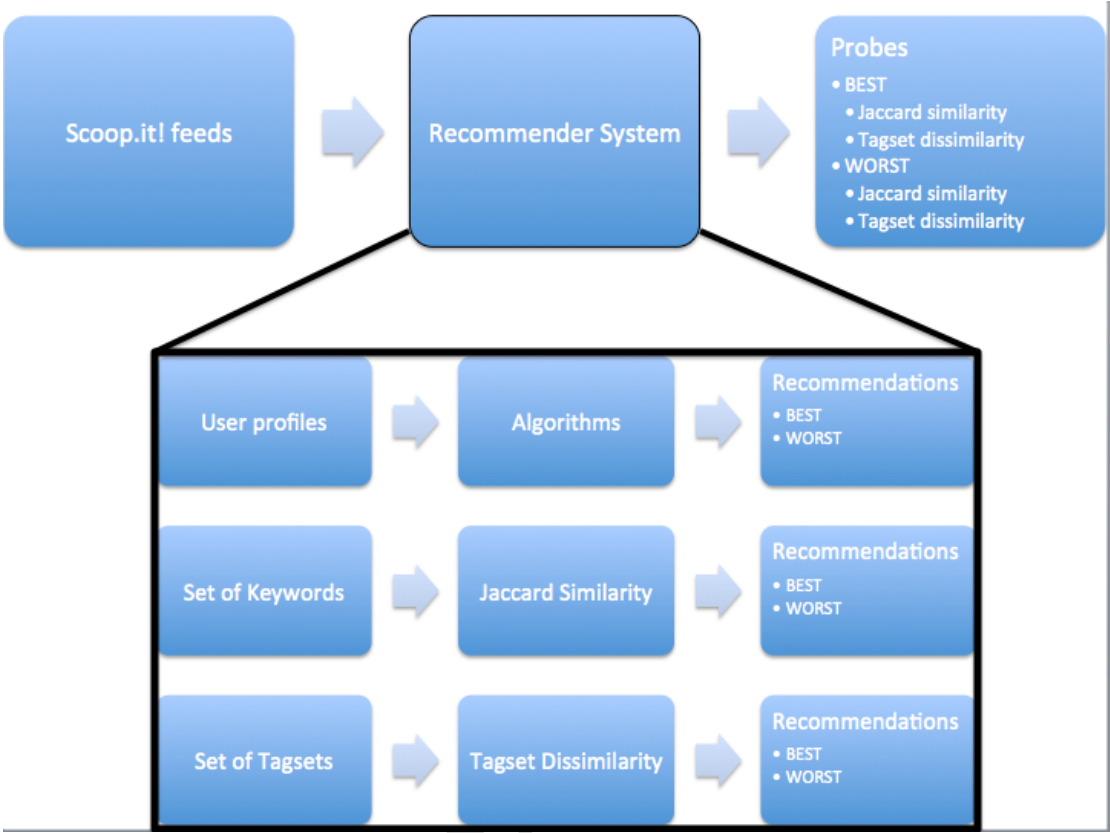
5.1.: Graphical illustration of the model described in Hypothesis 1



5.2.: Diagrammatic illustration of a Set of Concepts and a Set of Tagsets of a prototypical user



5.3.: Graphical illustration of the measuring instruments used in this experiment



5.4. Example of an original Scoop.it! post (top), the related Token-tagset (middle) and the retained Tagset (bottom)

Original Post

<name><name> XXXINSIGHTXXX literacy is the set of abilities and skills where aural, visual, and digital literacy overlap. These include the ability to understand the power of images and sounds, to recognize and use that power, to manipulate and transform digital media, to distribute them pervasively, and to easily adapt them to new forms.

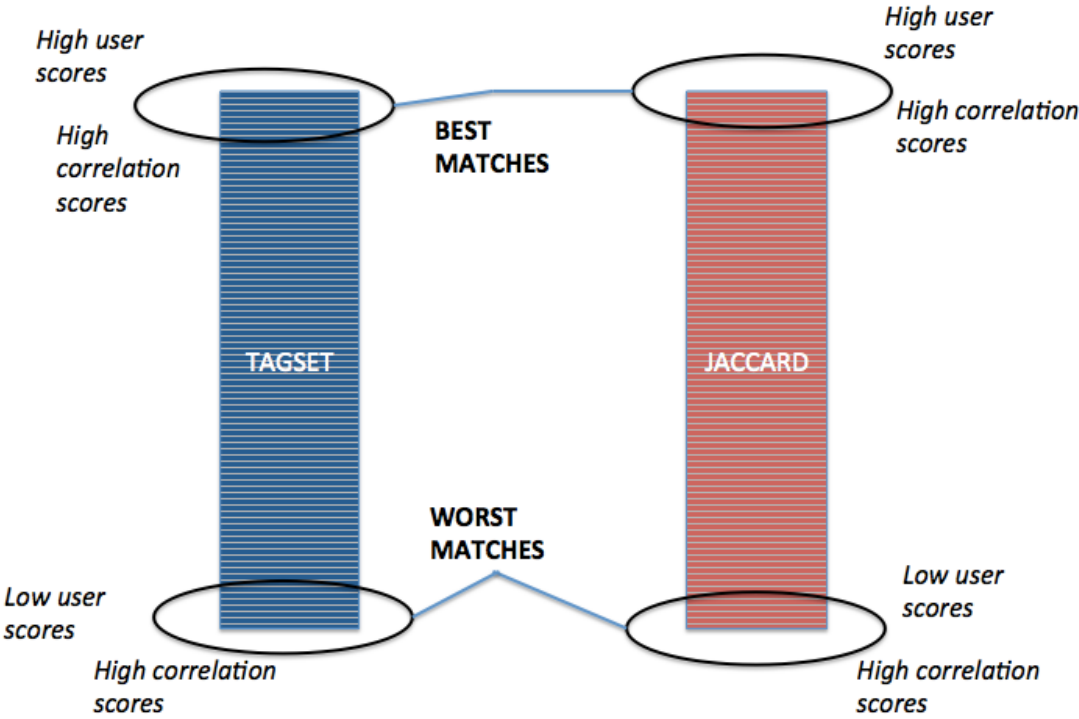
Token-tagset

['<name>', '<name>', 'XXXINSIGHTXXX', 'literacy', 'is', 'the', 'set', 'of', 'abilities', 'and', 'skills', 'where', 'aural', ' ', ' ', 'visual', ' ', 'and', 'digital', 'literacy', 'overlap', ' ', ' ', 'These', 'include', 'the', 'ability', 'to', 'understand', 'the', 'power', 'of', 'images', 'and', 'sounds', ' ', ' ', 'to', 'recognize', 'and', 'use', 'that', 'power', ' ', ' ', ' ', 'to', 'manipulate', 'and', 'transform', 'digital', 'media', ' ', ' ', 'to', 'distribute', 'them', 'pervasively', ' ', ' ', 'and', 'to', 'easily', 'adapt', 'them', 'to', 'new', 'forms', '.']

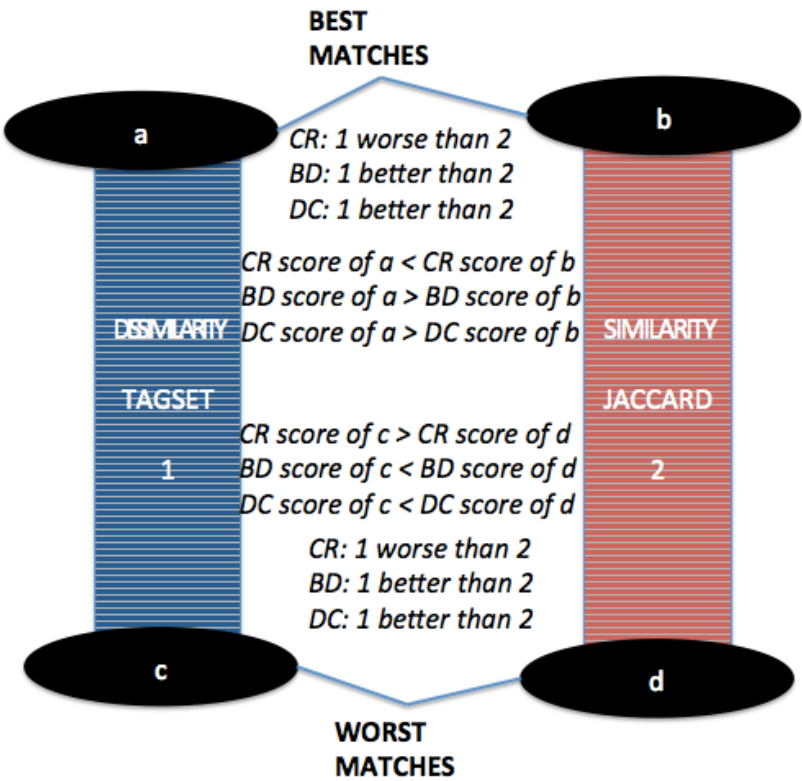
Tagset

[abil, adapt, easili, form, imag, includ, literaci, power, recogn, skill, sound, transform, understand, visual]

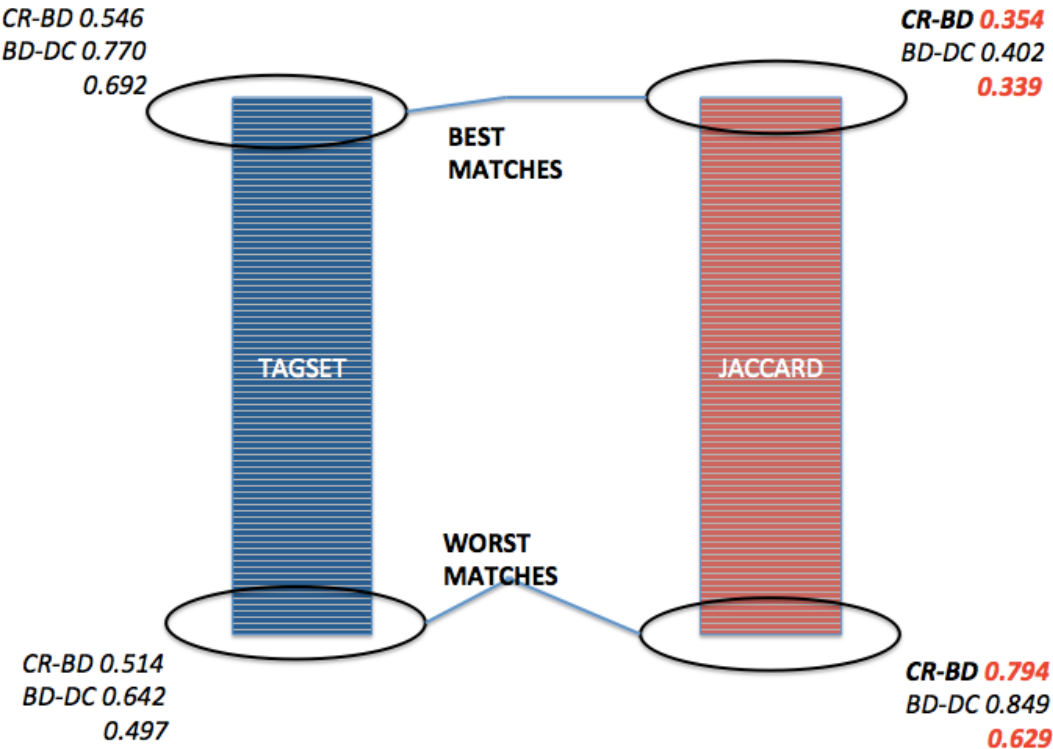
5.5: Hypothesis 1: Expected scores for *content relevance* (CR), *experience of breakdown* (BD) and the *desire to connect* (DC) and the expected correlation between these three



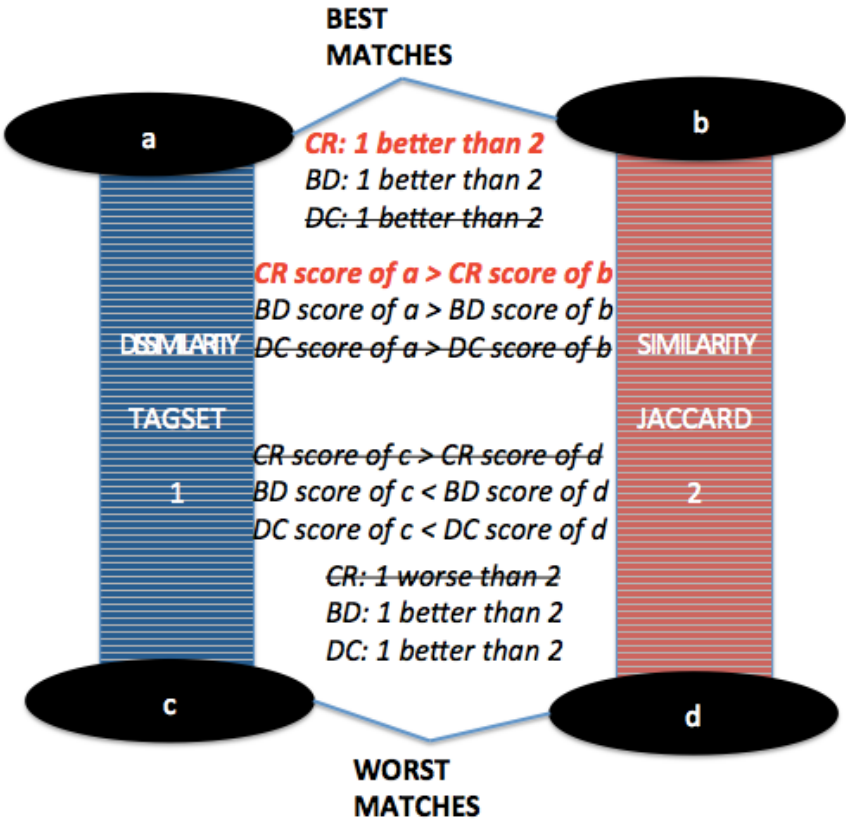
5.6: Hypothesis 2: Expected performance for two methods for *content relevance* (CR), *experience of breakdown* (BD) and the *desire to connect* (DC).



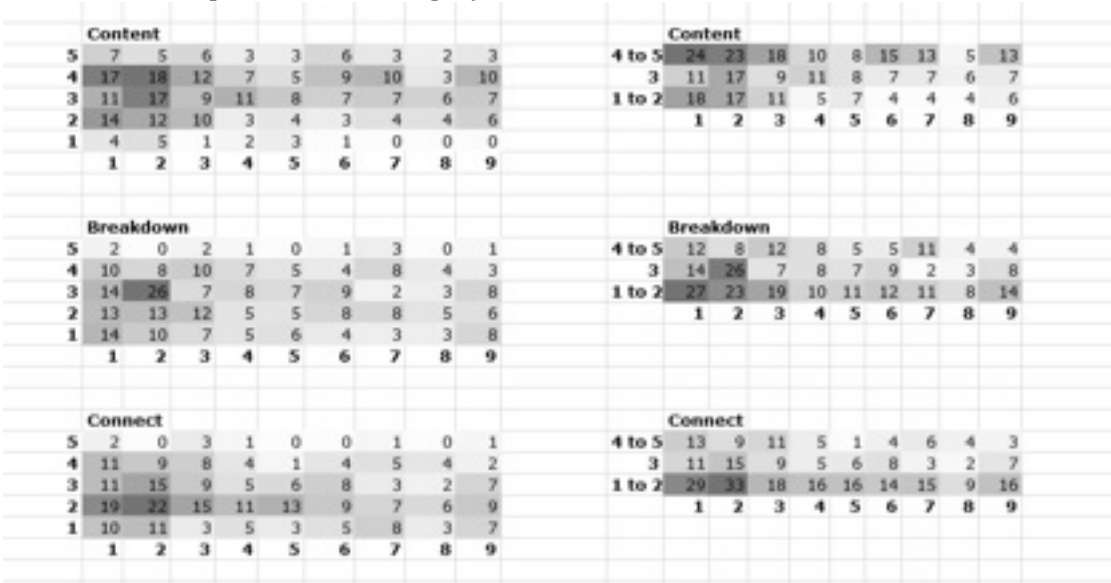
5.7: Hypothesis 1: Actual correlation scores for *content relevance* (CR) versus *experience of breakdown* (BD) and *experience of breakdown* (BD) versus *desire to connect* (DC): All results were significant. The results in bold did not match our initial expectations.



5.8: Hypothesis 2: Actual performance for two methods for *content relevance* (CR), *experience of breakdown* (BD) and the *desire to connect* (DC): the non-significant results have been stricken through. The results in bold did not match our initial expectations.



5.9: Heat maps of maximum intersection size vs. studied variables: *content relevance*, *experience of breakdown*, *desire to connect* (original 5-point scale on left, reduced 3-point scale on right)



Tables

5.1. Table 5.1 Processing steps of the user database to create user profile

Table 5.1 Processing steps of the user database to create user profile

	Process applied	Total number of pages	Total number of posts	Total number of unique units of meaning ('concepts')	Proportion of unique concepts per post
Step1	No process applied	240	32360	n.a.	n.a.
Step2	Tokenization	148	2957	15560	5.26
Step3	Removal of most frequent and least frequent words; stemming; removal of duplicates; grouping per curator	139	2886	2715	0.94

5.2. Table 5.2 Absolute and proportional numbers of matched tagset types

Table 5.2 Absolute and proportional numbers of matched tagset types

	Number	Percentage
Type A (full overlap)	19	0.00%
Type B (subset)	8386	0.11%
Type C (partial overlap)	1521503	20.55%
Type D (no overlap)	5874912	79.34%
Total	7404820	100%

5.3. *Table 5.3 Distribution of Jaccard scores and Tagset scores (n= 9593), $r=0.72$, $p<.01$*

Table 5.3 Distribution of Jaccard scores and Tagset scores (n= 9593), $r=0.72$, $p<.01$

	Minimum	Maximum	M	Variance	SD
Jaccard scores	0	0.501	0.0559	0.0037	0.0612
Tagset Scores	0	2.637	0.3683	0.1142	0.3379

5.4. *Table 5.4 Variables, their descriptions and the entities they belong to*

Table 5.4 Variables, their descriptions and the entities they belong to

Associated with entity	Variable	Description of variable
For each matched pair: For participant A with number of tagsets x and participant B with number of tagsets y	number of matched tagsets $k = x*y$	each matched pair of participants have a total number of matched tagsets, which is the product of both participants tagsets.
	number of tagsets per overlap type	Number of the total number of matched tagsets k in each type of overlap
	TypeA	TypeA - total overlap
	TypeB	TypeB - subset of
	TypeC	TypeC - partial overlap
For each participant who responded to the survey	TypeD	TypeD - no overlap
	number of tags in overlap in TypeC	In each of the matched tagsets pairs of type C with partial overlap, we checked the number of shared tags. This variable is the maximum of these numbers.
	Jaccard similarity score	Score for similarity for the matched pair
	Tagset dissimilarity score	Score for dissimilarity for the matched pair
	Evaluation Content Relevance for T1, T2, T3 and T4	Scores on a 5-point Likert scale, for each of the 4 recommended matches
	Evaluation Experience of Breakdown for T1, T2, T3 and T4	Scores on a 5-point Likert scale, for each of the 4 recommended matches
	Evaluation Desire to Connect for T1, T2, T3 and T4	Scores on a 5-point Likert scale, for each of the 4 recommended matches
	Preference to follow a recommended match for T1, T2, T3 and T4	Yes/no answer to if they want to follow a recommended match

5.5 *Table 5.5 Correlations between content relevance (CR) and experience of breakdown (BD)*

Table 5.5 Correlations between content relevance (CR) and experience of breakdown (BD)

Types of matches	Correlation between CR and BD
T1: Best dissimilarity	0.546**
T2: Best similarity	0.354*
T3: Worst dissimilarity	0.514**
T4: Worst similarity	0.794**
** p< .01, *p< .05	

5.6. *Table 5.6 Correlations and partial correlations between experience of breakdown (BD) and desire to connect (DC) (controlling for content relevance CR)*

Table 5.6 Correlations and partial correlations between experience of breakdown (BD) and desire to connect (DC) (controlling for content relevance CR)

Types of matches	Correlation between BD and DC	Partial correlation between BD and DC, controlling CR
T1: Best dissimilarity	0.770**	0.692**
T2: Best similarity	0.402**	0.339*
T3: Worst dissimilarity	0.642**	0.497**
T4: Worst similarity	0.849**	0.629**
** p< .01, *p< .05		

5.7. Table 5.7 Within-subjects Repeated Measures ANOVA results for user evaluations of variables content relevance (CR), experience of breakdown (BD) and desire to connect (DC) for best-ranked matches

Table 5.7 Within-subjects Repeated Measures ANOVA results for user evaluations of variables content relevance (CR), experience of breakdown (BD) and desire to connect (DC) for best-ranked matches

Variable	Wilks Lambda	F	p	Effect size	Power	Tagset dissimilarity		Jaccard similarity	
						M	SD	M	SD
Content relevance	0.922	3.784	0.058	0.078	0.478-0.5	3.52	1.13	3.15	0.965
Experience of breakdown	0.879	6.198	0.017	0.121	0.683-0.7	2.96*	1.30	2.46*	1.03
Desire to connect	0.948	2.477	0.123	0.052	0.338-0.34				

* significant difference in mean ($p < 0.1$)

5.8. Table 5.8 Within-subjects Repeated Measures ANOVA results for user evaluations of variables content relevance (CR), experience of breakdown (BD) and desire to connect (DC) for worst-ranked matches

Table 5.8 Within-subjects Repeated Measures ANOVA results for user evaluations of variables content relevance (CR), experience of breakdown (BD) and desire to connect (DC) for worst-ranked matches

Variable	Wilks Lambda	F	p	Effect size	Power	Tagset dissimilarity		Jaccard similarity	
						M	SD	M	SD
Content relevance	0.997	0.140	0.710	0.003	0.066				
Experience of breakdown	0.860	7.341	0.010	0.140	0.755	2.28*	1.26	2.72*	1.22
Desire to connect	0.866	6.943	0.011	0.134	0.732	2.28*	1.11	2.65*	1.25

* significant difference in mean ($p < 0.1$)